Translational Genomics Research Institute

http://www.tgen.org/

Company Location	*Contact Numbers*	*Contact 1*	*Contact 2*
400 N. 5th St., Suite 1600 Phoenix, AZ 85004	Tel 1: (602) 343-8400 Tel 2: (866) 370-8436 Fax: (480) 343-8440	Debbi Synder **Title:** Manager of Research Svcs. dsnyder@tgen.org (602) 343-8509	Galen P. Perry **Title:** Director of Outreach gperry@tgen.org (602) 343-8423

Overview

The ability to better diagnose, treat and ultimately cure disease in the 21st century will depend on two things: One, an improved understanding of the underlying genetics of disease and two, the ability and infrastructure to quickly translate this knowledge into diagnostics and therapeutics for patient benefit. The completed sequence of the human genome has profoundly shaped the way scientists now study and understand disease. This increased understanding will shift clinical practice from treatment based on symptoms, to treatment based on the underlying causes of disease in individuals. In turn, physicians will prescribe drugs that are more targeted in design, work more effectively, and have fewer side effects.

About TGen

The Translational Genomics Research Institute (TGen) is a private, non-profit research institute focused on accelerated translation of the sequenced human blueprint into therapeutics against human diseases. TGen uses broad-scale genome analysis tools and advanced bioinformatics capabilities to identify disease specific mutations and pathological biochemical pathways that are appropriate targets for novel therapy development.

Founded in June 2002, TGen is directed by Dr. Jeffrey Trent, the founding Scientific Director of the National Human Genome Research Institute at the National Institutes of Health (NIH), part of the Human Genome Project. Based in Phoenix, TGen is one of the best-equipped labs in the world for genomics research. Many of the investigators who worked at NIH with Dr. Trent have joined him as part of a major new biotech initiative for Arizona. As of May 2004, TGen had 150 employees and the number is projected to double to 300 over the next five years. Construction of the headquarters will be completed in November 2004.

TGen's Translational Research Strategy

At the core of TGen's strategy is the development of "accelerators" to quickly move research discoveries into the clinic to benefit patients. TGen is establishing these accelerators as vehicles to develop therapeutics, prognostics, and diagnostics and to deliver these discoveries to patients as quickly as possible. A description of three such accelerators follows:

1. The Center for Translational Drug Development (TD2)

Using animal models and genomic analyses, TD2 will provide development services to evaluate the safety and efficacy of pre-clinical compounds. It will also provide consultation and assistance with the initiation and coordination of clinical trials.

2. Molecular Profiling Institute

The Molecular Profiling Institute will provide testing to a nationwide base of patients using state-of-the-art genomic fingerprinting for diagnostic, prognostic, disease severity, patient response, and therapeutic targets.

3. Nanobiomics, Inc.

Nanobiomics, Inc. will target the rapidly growing molecular genetic diagnostic personal care market with integrated platforms and subsystem products and will enable drug development companies to more readily offer patients and physicians access to personalized medicine.

The accelerators are designed to not only feed TGen's clinical partners, but create a continuum to research and development efforts.

Research Divisions and Scientific Service Centers

Discovery fuels TGen's translational research and lies at the heart of our scientific investigations. Toward that end, TGen's research encompasses six divisions designed to foster a wide range of post-human genome discoveries. These research divisions are multidisciplinary in nature and draw heavily upon TGen's core technologies to expedite findings. Detailed below, they are as follows: Translational Drug Development, Pathogen Genomics, Genetic Basis of Human Disease, Molecular Diagnostics and Target Validation, Computational Biology, and Neurogenomics.

A. Translational Drug Development Division

To make individualized medicine a reality, discovery based research must be translated into tangible benefits for patients. We are applying the wealth of new data emerging from the human genome project and conducting research to determine why some patients benefit from treatment and others don't. Dr. Daniel Von Hoff, a world renowned scientist and clinician, directs the Translational Drug Development Division. His research laboratory is focused on genetic and biochemical approaches to identify molecular targets for the development of new therapies for disease, specifically pancreatic cancer. The team is involved in translating basic genetic, molecular, and biochemical observations and creating effective therapies for patients with pancreatic cancer. The objective is to develop targeted therapies based on distinguishing characteristic of pancreatic cancer cells versus normal cells. This approach relies on our ability to identify new targets in pancreatic cancer cells and agents that hit these targets. Once candidate genes (targets) are identified (as upregulated, amplified, deleted or mutated), a matching paradigm is used to identify agents which modulate those targets. As a result, TGen is poised to be a world leader in translating novel discoveries into individualized medicine for patients.

B. Pathogen Genomics Division

Pathogen Genomics is a joint program between TGen and Arizona's major universities, which results in a unified approach to protecting America against dangerous pathogens, in particular Bacillus anthracis (anthrax). Directed by Dr. Paul Keim, a leading expert on anthrax, the division encompasses work in the areas of comparative sequencing and computational biology. The division has three core missions:

1. Bolster the nation's bio-defense through improved forensic analysis;
2. Understand the interactions between man and microbe to develop new therapeutics and diagnostics to alleviate the human ailments caused by dangerous pathogens; and
3. Developing an improved understanding of disease movement to reduce and control the incidence of disease.

Building on the existing work at Northern Arizona University and the genomics tools at TGen, the Pathogen Genomics Division provides high-resolution genomic forensic analysis of biothreat pathogenic agents such as anthrax, plague, tularemia, ganders, brucellosis, and Q fever. These are the most dangerous of the bacterial bioterrorism and bioweapons agents.

The determination of the complete sequence of an organism represents the first step towards improved understanding of the complex interactions between man and microbe. By performing large-scale sequencing of entire pathogen genomes, such as plague, a complete genetic framework for intensive study of an organism provides insight into ways in which new and highly virulent pathogens evolve. The Program also will build upon our collaborations with state and federal agencies in developing molecular identification tools for public health related pathogens like E. coli, Salmonella, Listeria, and M. tuberculosis.

C. Genetic Basis of Human Disease Division

Armed with the knowledge that every disease has a genetic component, this division addresses important translational research questions associated with human disease. The division's researchers, directed by Dr. John Carpten, seek to understand the function of suspected disease-related genes and how they work in normal and "at risk" individuals, with particular emphasis on cancer and diabetes.

Accelerating this research are large-scale family studies centered on genetic samples collected from individuals within isolated populations. Information derived from these studies helps determine inheritance patterns for a genetic-related disease.

Identifying the patterns of disease occurrence in well-characterized families speeds the search for susceptibility or predisposition genes for that disease. Large-scale familial genomic approaches have recently been fostered in the area of multiple myeloma research. Additionally, the GBHD division has placed a strong emphasis in the area of type II diabetes, which aims to identify genomic factors involved in the development of progression diabetes and associated complications.

Through collaborations with other world-class investigators, GBHD scientists hope to contribute greatly to the overall knowledge of the genetics of human disease. The Divison encompasses four research units: melanoma, diabetes and obesity, genome variation and prostate cancer.

D. Melanoma Genetics Research Unit

Skin malignancies are the most common form of cancer in the US, and other parts of the world, including Australia. Dr. Pamela Pollock heads the Melanoma Genetics Research Unit, which focuses on characterizing the genetic events underlying the development of melanoma. The group utilizes genomic technologies to identify novel oncogenes and tumor suppressor genes that play a role in melanoma pathogenesis. In addition, the laboratory is characterizing several novel melanoma oncogenes to understand how aberrations of these genes lead to melanoma by studying the biochemical signaling pathways in which these genes function. Where applicable the unit will translate these genetic discoveries to advances in diagnosis, prognosis and treatment of melanoma.

E. Diabetes and Obesity Research Unit

At TGen, we recognize the threat of diabetes and the importance of combating this disease through education and scientific research. Our diabetes research team, led by Dr. Johanna Wolford, has the goal of identifying the causative or modifying genetic factors for the disease and its complications. The genes and molecular pathways we identify will form the basis for novel preventative and therapeutic strategies including methods of identifying high risk individuals, implementing intense interventions to prevent the development of the disease and its complications, and permitting early diagnosis.

F. Genome Variation Unit

Genetic variation plays a major role in risk associated with common complex diseases such as cancer, asthma, cardiovascular disease, and diabetes, to name a few. TGen hopes to make huge strides in the discovery of genetic variants that are associated with risk and development of disease. Dr. David Duggan heads the Genome Variation Unit, which utilizes cutting edge technologies and high throughput genomic analysis tools to facilitate the discovery of variants associated with disease risk. Dr. Duggan's current research is focusing around the identification of genetic variants associated with colon cancer risk.

G. Prostate Cancer Research Unit

Prostate Cancer is the most common cancer in US men. Led by Dr. Carpten, a team of researchers at TGen are among the most successful groups in identifying genetic risk factors associated with prostate cancer. Using positional candidate cloning approaches, Dr. Carpten's group is attempting it discover genetic variants which can be used for early detection of prostate cancer. Additionally, a large study is underway to identify risk factors associated with the disparity in prostate cancer seen in African American men. TGen hopes that these and other studies in the Prostate Cancer Research Unit will lead to tools, which will someone day make their way into clinical application.

H. Molecular Diagnostics and Target Validation (MDTV) Division

The Molecular Diagnostics and Target Validation (MDTV) division, led by Dr. Michael Bittner, is multidisciplinary in nature, drawing on biologists, computer scientists, engineers and mathematicians. Validation is critical for translational research to be realized. This team develops new tools and diagnostic methodology for validating potential targets.

TGen can be viewed as a discovery engine for the underlying causes of human diseases. As these causes are identified, technologies by which to accurately and precisely detect these markers become platforms for diagnostic tests. In order to fulfill its mission, TGen will validate genetic candidates of human disease and assess the performance of tests for detection of disease. This is already proving to be an area rich in the promise of novel technologies that will generate intellectual capital within TGen.

TGen is also establishing the robust and high-throughput platform of tissue microarrays (TMAs) for 28
accelerated validation of clinical markers of human disease. As candidate genes are identified and validated by TGen research teams (or in other institutes), new technologies advanced by TGen will enable probing studies of the functional consequences to a cell of turning-off these gene products. Such studies, conducted in very small volumes at a rapid pace, are pivotal to guiding subsequent commitments to optimize where in a biochemical cascade drug intervention would have the greatest benefit.

Several projects underway include the development of analytical tools to expand our understanding of the processes that cause and sustain cancer and the development of a mathematically sound way to determine whether coordinated behavior of small sets of genes in cells is statistically significant. TGen acquired unique resources to support functional chemo-genomics including large siRNA libraries, and chemical compound libraries for cell based screening. The development of RNAi microarrays as a new platform technology for miniaturized genetic silencing to enable functional genomics for the development of new and more effective anti-cancer drugs.

I. Computational Biology Division

TGen's Computational Biology Division conducts research in several areas including high performance biomedical computing, knowledge based scientific data management, Boolean network modeling,

biostatistics and bioinformatics. A clinical information management system has been developed, as well as software for web based analysis and visualization for a microarray consortium that the neurogenomics division leads. In addition, several documents and tutorials in biostatistics and bioinformatics have been created and disseminated for the purposes of analyzing genomic research data.

IBM in a partnership with TGen and Arizona State University (ASU), designed a supercomputer architecture to cost-effectively scale to thousands of processors for high-throughput computing. The result is one of the most powerful systems in the world. The ASU/TGen High Performance Computing Center provides computational resources, biomedical informatics support, and data management systems to investigators so that they might apply innovative information technology to biomedical research problems. The supercomputer ranks among the top 50 most powerful computers in the world.

TGen's high-performance scientific computing facility eliminates bottlenecks and provides researchers reliable access to a variety of resources including:

1. High performance computational tools such as Sequence Alignment, Linkage Analysis, Gene Clustering and Classification, and Multivariate Analysis.
2. Data management systems such as Sequence, SNP and Expression DB, as well as Gene Ontology.
3. Customized application software such as Genomic Regulatory Network Simulations.

J. Neurogenomics Division

The Neurogenomics Division, led by Dr. Dietrich Stephan, seeks to develop improved diagnostics and therapeutics for neurological and mental health disorders using tool sets and well established patient cohorts for clinical analysis. Four research units within the division focus on a particular set of diseases which form the basis for its experimental efforts. Additionally, Dr. Eric Reiman, Neurogenomics Clinical Director, is an internationally recognized researcher in the study of Alzheimer's disease (AD), and the director of a nationally recognized statewide Arizona consortium that is poised to leverage TGen discoveries into the clinical evaluation of putative drug therapies for the treatment and prevention of AD and other neurological and psychiatric disorders. Some of the diseases studied by each of the Division's active Research Units are as follows:

a. Neurodegenerative Unit

The Neurodegenerative Unit studies diseases such as Alzheimer's disease, Amyotrophic Lateral Sclerosis (ALS or Lou Gehrig's disease), Parkinson's disease, and Multiple Sclerosis. The Unit utilizes brain imaging techniques that have revolutionized the diagnosis of Alzheimer's disease, and this technology, as part of a collaboration with three Alzheimer's Disease Centers (in Arizona, Washington University and Duke University), will soon be providing new information about the molecular processes involved in this devastating disorder. The National Institutes of Health awarded to TGen one of only three national Alzheimer's disease Genomics Center awards. This grant capitalizes on TGen's partnership with the Sun Health Research Institute and the Arizona Alzheimer's Disease Consortium. The Neurodegenerative Unit also has a large study ongoing with Harvard University which has recently identified two new regions of the genome containing ALS-causing mutations.

b. Brain Tumor Unit

The Brain Tumor Unit focuses on the identification, development, and clinical testing of new diagnostics and therapeutics against adult and pediatric primary brain tumors. The unit focuses on brain cancers such as adult glioma, pediatric astrocytoma, and pediatric medullablastoma. Primary research partners include the Barrow Neurological Institute, Arizona State University, and the Mayo Clinic Scottsdale.

c. Neurobehavioral Unit

The diseases studied by TGen's Neurobehavioral Unit include schizophrenia and autism, the latter of which is the subject of an exciting collaborative effort in partnership with the Southwest Autism Research and Resource Center (SARRC) to conduct the most comprehensive genetic and molecular study of autism to date. The goal of the SARRC/TGen collaboration is to identify the underlying pathogenesis of autism through the use of advanced tools for the analysis of genomics and proteomics. This approach will compare cases with appropriately matched controls in order to identify one or many genes, RNA transcripts, and/or proteins that are highly correlated with the autistic phenotype. The unique aspect of the research will involve population-based genome-scanning, an approach that has never been conducted before in autism due to technological constraints.

d. Neurodevelopmental Unit

TGen's Neurodevelopmental Unit studies diseases such as Niemann-Pick Type C, Swarey Syndrome, and Huntington's disease. Dr. Stephan's research team at the National Human Genome Research Institute identified the genetic mutation responsible for Swarey Syndrome. This was accomplished, from start to finish, in 5 days – the fastest time ever for the identification of a disease-causing gene. The Neurodevelopmental Unit is collaborating with the Amish community in Lancaster, PA, which experiences an unusually high incidence of the disease.

Scientific Service Centers

The development and application of modern genome technology will be an essential component of TGen's objective of performing contemporary genetic analysis. Researchers in this area study the structure and function of genomes in normal and disease states, and have expertise in a wide range of genomic techniques. These techniques support TGen's internal scientific programs by facilitating the translation of data into practical applications. TGen's Scientific Service Centers include RNAi, Gene Expression Microarray, Tissue Microarray, DNA Sequencing, Bioinformatics, and SNP Genotyping.

A. RNAi Service Center

The Cancer Drug Development Laboratory (CDDL) provides RNAi services to all TGen investigators. RNAi technology has the potential to enable and accelerate a wide range of basic research, including large-scale functional genomics for drug target identification and validation, as well as functional chemogenomics to discover and validate genes that modulate cellular response to chemotherapy.

Dr. Spyro Mousses, the laboratory's director, has extensive experience in the field of functional genomics and significant expertise in the development and application of RNAi technology. The CDDL is pioneering the development of extensive technologies, infrastructure, and RNAi resources to conduct genome scale functional analysis in human cells for numerous molecular endpoints and cellular phenotypes, including response to chemotherapy. The CDDL aims to integrate these capabilities into systems biology approaches to accelerate the drug discovery process and effectively translate genomic data rapidly to new anticancer therapeutics.

The CDDL's three-tier approach to accelerating drug discovery involves 1) the integration of data from several different microarray platforms to better understand the molecular mechanisms that play a key role in cancer and to prioritize the prime targets for drug development, 2) development and application of innovative high throughput RNAi based technologies allow genome scale drug target discovery and validation, 3) the development and discovery of small molecule inhibitors, monoclonal antibody, and nucleic acid based agents to target the most promising gene products for advancement to preclinical and

clinical testing. This research is setting the stage for accelerating the translation of genomics for the development of new and more effective anti-cancer drugs.

B. Gene Expression Microarray Service Center

Lead by Dr. Dietrich Stephan the Gene Expression Microarray Center offers experimental design advisory services, expression profiling array analysis, DNA analysis array (hybridization and scanning), and data analysis, including visualization and statistical analysis. Together with the significant infrastructure and throughput available at TGen, investigators and their collaborators seek to rapidly garner validated expression fingerprints of pathogenic processes.

Dr. Stephan is the Chairman of a national consortium of neurological expression profiling centers (NINDS-NIMH Array Consortium) that services a user base of ~5,000 extramural investigators. One arm of the consortium, housed within TGen, indicates our ability to manage high-throughput projects and the high quality of generated data, all of which is NIH sanctioned.

C. Tissue Microarray Service Center

TGen's Tissue Microarray (TMA) Center provides a valuable tool for the validation of molecular markers in clinical specimens. Directed by Dr. Galen Hostetter, a board-certified pathologist with extensive experience in TMA design, construction, and analysis, the Center adheres to the highest level of quality control, thus maximizing extraction of biological information from archival tissue blocks.

The TMA plays a critical role in the translational research process through standardized collection and organization of clinical tissues. The TMA format offers high throughput capacity for molecular analyses and provides inherent assay standardization. Active collaborations with imaging experts and computer programmers are an integral component of the Center's efforts to continually extend the applications of the technology. Most notable of these was a two-year collaborative development of an automated fluorescence in situ hybridization (FISH) spot counting program. By design, this highly interactive program enumerated probe signals after hybridization of specific DNA sequences in tumor specimens using the TMA format. In situ hybridization (ISH) applications using DNA or RNA probes remains an active part of Dr. Hostetter's research.

D. DNA Sequencing Service Center

TGen's modern, fully automated DNA sequencing facility leverages today's technology to address urgent biomedical questions. Additionally, as the research appetite for sequence information increases, it is of broad scientific interest to continue to develop and implement new methods and technologies for DNA sequencing.

TGen's Sequencing Center, directed by Dr. Jeffrey Touchman, covers many areas of experimental, theoretical, and computational genomics, and links to areas like functional genomics and molecular evolution as well as interdisciplinary connections to statisticians and computer scientists. These foundational services provide the springboard from which TGen research can move forward into many disease areas.

The technology is powerful enough to sequence an entire genome, but also sensitive enough to identify irregularities, or mutations, in the sequence of individual genes. The Center houses a high-throughput, state-of-the-art lab employing two of the most advanced DNA sequencing platforms available today and is supported by robotic automation to ensure high quality, unparalleled accuracy, and fast sample turnaround.

The DNA Sequencing Center offers high-throughput single pass sequencing of plasmid and PCR DNA templates; Shotgun sequencing (draft or high-accuracy) of BAC, PAC and Cosmid clones; BAC-end sequencing; Shotgun DNA library construction; Full-length cDNA sequencing; and Automated clone picking/arraying. Sequencing data generated by Center provides scientists with information about new genetic markers, new diagnostic capabilities, and a greater understanding of the mechanisms of disease.
Bioinformatics Service Center

Under the leadership of Dr. Ed Suh, the Computational Biology and Bioinformatics (CCB) Center engages in computational bioscience research and development collaborations with the TGen biomedical research programs; and provides computational resources, biomedical informatics support, and knowledge based data management systems to the TGen scientific staff. CCB conducts its research and development program with five principal objectives:

- To develop and apply biostatistics, pattern recognition, data mining, text mining, and data visualization techniques to complex biomedical and clinical research problems.
- To develop and provide knowledge-based data management systems for the discovery of biomedical knowledge, including clinical informatics and management systems to develop and provide high performance computational methods and algorithms to analyze large volumes of scientific data and to simulate complex biological systems.
- To provide high performance computing resources and scientific research tools, including special-purpose parallel computing machines and high-density storage systems.
- To collaborate with the TGen scientific staff and colleagues at other research centers in applying state-of-the-art information technology to biomedical research problems.

The Center staff are experts at manipulating and making sense of complex data sets. This team has been responsible for developing some of the most sophisticated algorithms for understanding gene-gene interactions and complex signatures that correlate to the phenotype of interest.

TGen's state of the art bioinformatics and computational biology facility lends support to all disease related program studies. This group develops custom solutions to complex bioinformatics problems. It is equipped to provide bioinformatics support, consultation services, and data analysis for a wide array of research efforts. By employing large datasets, TGen's powerful scientific computing environment, ranked in the top 60 worldwide in terms of computing power, provides researchers with ways to model more realistic biological systems, discover complex patterns, and determine multivariate associations among data elements.

E. SNP Genotyping

Single Nucleotide Polymorphisms (SNPs) are the most abundant form of genetic variation in the human genome, with an approximate frequency of one every kilobase. SNPs are likely to give us the greatest power in identifying disease causing genes, associations, and DNA fingerprints, all of which have profound implications for detection, prognosis, and therapeutic intervention. Until recently however, efficient, accurate, and cost-effective SNP genotyping technologies were a rate-limiting step in the exploitation of these powerful forms of genetic variation.

TGen's High-Throughput SNP Genotyping Center, directed by Dr. David Duggan, addresses this bottleneck. By strategically aligning itself with the leading providers of high-throughput SNP genotyping technology, the Center offers the ultimate genotyping solution that is both integrated, meets demands of genotyping programs, and addresses the requirements for flexibility, simplicity, throughput, economy, and accuracy. The combination of these latest technologies provides access to the complete gamut of

SNP genotyping services, including: whole genome analysis (linkage or association), fine-mapping, and candidate gene studies.

Microsatellite-based linkage analysis is also provided via the Centers microsatellite linkage panel (LMP v2.5, Applied Biosystems). The Center is complemented by and collaborates with other groups within TGen. For example, the Center works closely with TGen's DNA Sequencing Facility on SNP discovery. By making use of a variety of control samples, we can discover and validate SNPs for poorly characterized regions of the genome, specific genes, and/or pathways. The Center also integrates with TGen's Statistical Group, and has access to high performance computing resources through TGen's Bioinformatics / Computational Biology Core. By utilizing these resources, the Center can bring greater power to bear on its analyses, and provide a more complete range of services.